

COLOR FEATURES FOR DATING HISTORICAL COLOR IMAGES

Basura Fernando^{†}, Damien Muselet^{*}, Rahat Khan[†] and Tinne Tuytelaars^{*†}*

^{*†} PSI-VISICS, KU Leuven, iMinds, Belgium

^{*} Universit Jean Monnet, LaHC, Saint-Etienne, France

[†] ALCoV-ISIT, UMR 6284 CNRS / University of Auvergne, Clermont-Ferrand, France

ABSTRACT

Estimating the age of historical photographs is a challenging task for human beings. Only recently this task has been addressed in computational image analysis perspective. The characteristics of the device used to acquire each photograph are discriminative features for this task. We aim at extracting such characteristics from a historical color photographs. The acquisition device mainly effects two properties of the colors: the distribution of their derivatives and the angles drawn by three consecutive pixels in the RGB space. We propose two color features that take advantage of these observations. We show that these two color descriptors (namely color derivatives and color angles) attain the state-of-the-art in the context of image dating.

Index Terms— Dating images, Color features, Color derivatives, Color angles, Imaging process

1. INTRODUCTION

The age of an old photograph is a useful information for various areas of studies such as history, genealogy and anthropology. They are interested in the conservation and preservation of old photographs. More often than not this information is not readily available. Moreover, there is a growing interest towards cultural heritage projects such as Flickr Commons [1] that tries to make historical photography collections accessible online. These preservation projects have a real necessity to estimate the age of a photograph automatically. Only recently, the task of automatically estimating the age of historical photographs has gained some attention [2, 3].

Despite its limitations, manually estimating the age of the photograph seems to be the most popular method at the moment. Mostly, these manual techniques rely on visual cues such as particular fashion or hair styles worn by human subjects [4, 5]. These approaches are not reliable and can't be applied to all kinds of photographs. In some specific instances, estimating the age of the photograph can be done using domain specific knowledge such as in [6] where military historical photographs are successfully dated using such techniques.

Following the work of Palermo et al. [2], we propose to solve the task of classifying historical color photographs into



Fig. 1. Some photographs from left to right in chronological order. Left photograph is from 1930s next ones from 1940s, 1950s, 1960s and the last from 1970s. This figure clearly shows the difficulty of the task.

decades (see Fig. 1). In this aim, Palermo et al. propose to analyze the variation of color distributions across the age to classify the images. In particular, they exploit the probability to have a certain color saturation for each considered hue. The hue and saturation they are using are estimated from the CIELAB color space although the transformation from the available RGB components to this standardized color space is not accurate in such uncalibrated acquisition conditions. Furthermore, we show that, given one image, while being very useful, its color distribution is not the most accurate feature to discriminate between decades.

Unlike classical color descriptors, our proposed features are device-dependent. We consider that the characteristics of the devices used to acquire one photograph are very accurate features to date this photograph. For example, the color films used in 1930s are different from the once used in 1970s. Therefore, we aim at extracting these characteristics from each image. From the imaging process, we claim that the acquisition devices impact two different properties of the colors: first, the global distributions of the RGB derivatives and second, the shapes drawn in the RGB space by consecutive pixels along the edges. Thus, we propose to use two different and complementary color features for the dating task, namely **color derivatives** and **color angles** introduced in section 3. We show in section 4 that they perform significantly better than the state-of-the-art [2] and than novel features such as DeCAF [7] which are based on deep learning [8].

2. RELATED WORK

The most related work for dating historical images is the one from Palermo et al. [2]. They propose to split the available historical images into discrete time interval classes and define the problem as an image classification problem. While many definitions of these time intervals are possible, decade (e.g. the 1950s) based grouping is intuitive from cultural trends point of view. A novel color descriptor called *Conditional Probability of Saturation Given Hue* is proposed. This descriptor is based on the argument that unique appearance of images produced by historical color imaging process is at least partly due to differences in their reproduction of certain hues, especially with regard to saturation [9].

Schindler et al. [10, 11] proposed a method to temporally order a large collection of historical skyscraper images by reconstructing the 3D world using structure-from-motion which requires many overlapping images of the same scene over a large time lag. Following a similar idea Xu et al. [12] proposed a method to model 3D urban scenes in the spatial-temporal space using a collection of photographs that span over many years. Recently, Rematas et al. [13] try to infer the age of a man made object in an image using modern descriptors such as SIFT and encoding methods such as Fisher vectors, while Lee et al. [3] try to discover specific styles using mid-level features. These methods do not inspect the photograph as a physical artifact but rather rely on object recognition techniques to identify specific patterns that are correlated with time or the decade. Even though, these methods can be useful, they are limited by the image content of the historical photograph. That's why in this paper, we propose an approach that is not restricted to some kind of contents and that is based on two different and complementary color descriptors introduced in the next section.

3. PROPOSED COLOR DESCRIPTORS

We also define the task of estimating the age of a photo as determining the decade during which a historical color photograph was taken. In this section we introduce two color descriptors that are useful to classify images based on the age of the photograph. The intuition behind these descriptors is as follows: the color content of the images have changed over the years because of the evolution of the acquisition devices. By analyzing the photograph contents, we try to extract information that reveals the camera characteristics and then use them to infer the age of the photograph.

3.1. Color acquisition process

From [14], it is clear that different cameras do not exhibit consistent color responses while observing the same scene under the exact same conditions (viewpoint, illumination, ...). Indeed, each acquisition device uses its own mapping process

from scene radiance to pixel intensity. As explained in [15], each step of this acquisition process can be associated with one specific (and often unknown) transform. In this paper, we claim that some of these transforms are impacting the global statistics of the colors in the resulting photographs while others make the relation between scene radiance and pixel intensity nonlinear. For example, the white balance and the color transform are changing the color statistics whereas tone mapping, aesthetic effects, gamma correction or old film responses are adding non-linearity to the mapping [16].

Consequently, in order to measure the impact of the acquisition systems (changing over the years) on colors, we propose two color features that are related to color statistics and non-linearity assessment.

3.2. Color derivatives

Palermo et al. [2] use color statistics to evaluate the age of a photograph. They approximate the $L^*a^*b^*$ components from the available RGB, which is a coarse approximation in such uncalibrated acquisition conditions. Then they deduce hue and saturation from these components. In this paper, instead of using approximate transforms between color spaces, we propose to exploit the variations of the color distributions over the time directly in the original RGB camera color spaces. Our motivation comes from the fact that RGB global color distributions have been shown to reveal information about the acquisition conditions such as the lights and the camera [17, 18]. That's why, in most color constancy algorithms, the analysis of such distributions helps in removing the impact of the acquisition conditions on the colors. For example, the average RGB components are used by the grey-world algorithm and the convex-hull of the RGB distribution is the basis of the gamut-mapping algorithms [19]. More recently, it has been shown that taking into account the 3D distributions of the first and second derivatives of the RGB components improves the results of the classical algorithms [18, 20]. The idea behind these algorithms is that RGB zeroth, first and second derivatives coarsely follow canonical distributions in every natural image and deviate from these distributions because of the lights and cameras used during the acquisition. Thus, we claim that analyzing these distributions can help in extracting information about the acquisition conditions and thus in estimating the age of a photograph. Therefore, we propose to extract the RGB components, the RGB first and second derivatives and to construct one 3D histogram for each.

3.3. Color angles

Whatever their age, all cameras exhibit a non-linear relationship between the scene radiance and the pixel intensity. Nevertheless, these non linear transforms are different between cameras. The idea of camera radiometric calibration is to estimate such transform for a specific camera [21]. In this paper,

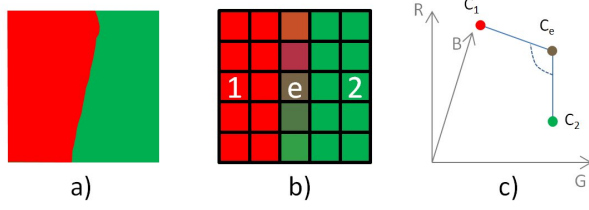


Fig. 2. Color angle illustration inspired from [2]. a) Two homogeneous regions in a scene. b) The projection of these colors on the image plane. c) The corresponding colors C_1 , C_2 and C_e are not aligned because of the non-linear transform and the angle is specific to the used acquisition system.

we exploit these device-dependent features and propose to estimate and use this non linear transform to characterize the camera that acquired each considered photograph. In this aim, we exploit the work of Lin et al. [2] who analyze the RGB positions of the colors along the edges in the image. Indeed, let us consider the border between two homogeneous regions in a scene observed by an imaginary camera whose mapping function (radiance to intensity) would be linear. During the acquisition, the projection of these regions in the image plane forms two homogeneous regions whose colors are respectively $[R_1, G_1, B_1]^T$ and $[R_2, G_2, B_2]^T$ in the image. Furthermore, the colors $[R_e, G_e, B_e]^T$ of the edge-pixels which lie along the border between these regions are linear combinations of these two colors. Consequently their colors should be on the line connecting the two colors $[R_1, G_1, B_1]^T$ and $[R_2, G_2, B_2]^T$:

$$\begin{aligned} \exists \alpha \in [0; 1] \text{ so that,} \\ [R_e, G_e, B_e]^T = \alpha [R_1, G_1, B_1]^T + (1 - \alpha) [R_2, G_2, B_2]^T. \end{aligned} \quad (1)$$

However, since in real cases, the camera mapping f from scene radiance to pixel intensity is not linear, the three colors $f([R_1, G_1, B_1]^T)$, $f([R_e, G_e, B_e]^T)$ and $f([R_2, G_2, B_2]^T)$ do not form a line in the RGB space, i.e. eq. 1 does not hold any more. Let us denote C_1 , C_2 and C_e the three considered colors. In the imaginary case of linear mapping, the angle $\widehat{C_1 C_e C_2}$ is flat ($= \pm\pi$), while in real cases, its value is different from $\pm\pi$. This point is illustrated in fig. 2. Even, for a given camera, this angle is not constant and depends on the colors C_1 , C_2 and C_e .

For clarity, above, we have considered the 3D space RGB where one single angle can be estimated for each edge-pixel. But in order to get more information about the non linearity of the camera, we can project the 3 colors C_1 , C_2 and C_e on each plane RG , RB and GB and measure one angle in each plane. We have experimentally validated that accounting 3 angles a_{RG} , a_{RB} and a_{GB} increases the discriminative power of our color feature with respect to a single "3D" angle a_{RGB} .

Therefore, we propose to analyze the angles formed in

the RG , RB and GB planes by all triplets of consecutive (in the image space) pixels (in both horizontal and vertical directions). Since the values of these angles vary according to the components R_e , G_e and B_e of the considered edge-pixel, we propose to measure the evolution of this angle across these components.

Consequently, for each pixel triplet associated with the edge-pixel color $[R_e, G_e, B_e]^T$, we evaluate 3 angles a_{RG} , a_{RB} and a_{GB} and accumulate this information for all the triplets of one image in 6 co-occurrence matrices that count the number of times each color angle occurs with each related color component $M_{a_{RG}}^{Re}$, $M_{a_{RB}}^{Re}$, $M_{a_{RG}}^{Ge}$, $M_{a_{GB}}^{Ge}$, $M_{a_{RB}}^{Be}$ and $M_{a_{GB}}^{Be}$. All the triplets are considered without edge selection.

4. EXPERIMENTS

4.1. Decade Classification

To evaluate the effectiveness of our proposed features, we use the same dataset as [2]. This dataset is composed by decades from 1930's to 1970's (five classes see Fig 1). It contains 1375 images in total. As done in [2], we use 225 images per class for training and the rest for testing. We use both linear and non-linear (chi-square kernel) SVM classifiers with a fixed cost parameter ($C=100$ in LibSVM). We report classification accuracy averaged over ten random training/testing splits for each feature.

We test the performance of the proposed color features. First, the color derivative features described in section 3.2 consist of 3D histograms of RGB components (0-CD), first derivatives (1-CD) and second derivatives (2-CD). Each histogram is 10^3 dimensions (10 bins per channel) and is tested independently from the others. The combination of these three color statistics features is also tested (3000-D). Second, the color angles (CA detailed in section 3.3) consist of 6 co-occurrence matrices, each one being 5^2 dimensions (5 bins for angles and 5 bins per channel). Thus, this descriptor is 150 dimensional. These color features are compared with:

CIELAB Histogram : This is a color feature commonly used in previous works such as scene recognition [22] and image geo-location [23]. As proposed in [2], this descriptor is $5 * 15 * 15$ (1125) dimensions.

P(Sat|Hue) descriptor: This descriptor is introduced in [2]. It is based on the argument that the unique appearance of images produced by historical color imaging process is at least partly due to differences in their reproduction of certain hues, especially with regard to saturation. This descriptor is 512-D.

DeCAF descriptors: A deep convolution model is trained in a fully supervised setting using a state-of-the-art method [8]. Then various features are extracted from this deep network. The activation of n hidden layer of the deep convolution neural network is denoted by $DeCAF_n$. These activation values are used as features.

Table 1. Experimental results using both linear and non-linear SVM classifiers.

Method	Accuracy %	
	Linear	Non-linear
Human	26.0 [2]	
Best results from [2]	45.7	
CIELAB Histogram	37.3 [2]	43.2 \pm 3.0
P(Sat Hue) descriptor	37.6 [2]	43.1 \pm 2.5
<i>DeCAF</i> ₆	46.7 \pm 4.7	18.4 \pm 2.9
<i>DeCAF</i> ₇	42.8 \pm 2.1	18.8 \pm 3.0
Our features		
RGB Histogram (0-CD)	41.8 \pm 1.8	47.1 \pm 3.6
RGB First derivatives (1-CD)	45.5 \pm 1.6	47.6 \pm 1.7
RGB Second derivatives (2-CD)	47.8 \pm 2.1	50.5 \pm 2.6
Color angles (CA)	54.4 \pm 4.0	55.4 \pm 3.3
0-CD + 1-CD + 2-CD	45.9 \pm 3.3	84.3 \pm 2.1
0-CD + 1-CD + 2-CD + CA	48.2 \pm 3.2	85.5 \pm 1.5

We apply L2 normalization for all features. We report results using both linear and non-linear (chi-square) kernels for each feature in Table 1.

First, we analyze the results provided by linear SVM (left column of Table 1). We notice in these results that human observers find this task extremely difficult (accuracy of 26.0%). On the other hand, it is interesting to see that color-based approaches perform better than human which is rare in image classification tasks as reported in [2]. The CIELAB histogram reports a classification accuracy of only 37.3% in [2] while using the RGB histogram we obtain 41.8%. This confirms our reasoning for using RGB data directly for this task. In [2], the best results provided by a single feature are those of the conditional probability of the saturation given the hue ($P(\text{Sat}|\text{Hue})$) which reports 37.6%. All our color statistics features (0-CD, 1-CD and 2-CD) outperform this accuracy. This clearly shows that the RGB derivative distributions capture useful information about the acquisition device hence the information about the age of the photograph. The color angle feature (CA) performs the best as an individual feature with a significant performance of **54.4%** despite its relative small dimensionality. Indeed, the best result reported in [2] is equal to 45.7% and corresponds to a fusion of many (7) classical features. Finally, modern state of the art object recognition features such as *DeCAF*₆ or *DeCAF*₇ which are based on the deep learning do not perform as good as hand crafted features such as the color angles for this specific task.

By analyzing the distributions of all the histogram features, we noticed that some bins are over represented because some colors, derivatives or angles are very frequent over all the images. That specificity of the histograms explains why we also propose to test chi-square SVM which normalizes the difference between two histogram bin values by the sum of these values. Obviously, we notice that all the

Confusion Matrix

	1	2	3	4	5	
1	379 19.0%	11 0.6%	6 0.3%	4 0.2%	9 0.5%	92.7% 7.3%
2	9 0.5%	326 16.4%	24 1.2%	7 0.4%	25 1.3%	83.4% 16.6%
3	2 0.1%	15 0.8%	340 17.1%	16 0.8%	21 1.1%	86.3% 13.7%
4	3 0.2%	18 0.9%	7 0.4%	338 17.0%	27 1.4%	86.0% 14.0%
5	7 0.4%	20 1.0%	23 1.2%	35 1.8%	318 16.0%	78.9% 21.1%
	94.8% 5.2%	83.6% 16.4%	85.0% 15.0%	84.5% 15.5%	79.5% 20.5%	85.5% 14.5%
	1	2	3	4	5	
	Target Class					

Fig. 3. The confusion matrix summed over 10 random splits for the best performing approach (non-linear SVM with 0-CD + 1-CD + 2-CD + CA). Note that the classes [1, 2 ..., 5] correspond to [1930s, 1940s, ..., 1970s].

histogram based features improve when moving from linear to non-linear SVM. This is more outspoken when dealing with feature fusion which seems to be badly adapted to linear kernels.

All the previous comments remain the same with non-linear SVM, i.e. our descriptors also outperform the state-of-the-art with non-linear SVM. Finally, by late fusing *RGB histograms*, *1-CD*, *2-CD* and *CA* features we manage to obtain the very high accuracy of **85.5%**. So our descriptors work three times better than an average human and twice as good as the state of the art. This result clearly shows that our color features are highly complementary.

In Fig. 3, we show the cumulative confusion matrix computed over ten random splits when using the fusion of all our proposed descriptors with non-linear kernel. The best results are obtained for 1930s (94.8%) while the worst results are obtained for 1970s (79.5%). This is not surprising given the fact that during 1930 era there were only a few number of color films available.

5. CONCLUSION

In this paper we present a successful method to date historical color photographs. By utilizing RGB derivatives and color angle features, we collect a significant amount of information about the photograph capturing process. These features are used to classify old color photographs into decades. While an average human performs at a classification rate of 26%, we manage to obtain a good performance of 85.5% using the proposed color features and even outperforming state of the art deep learning features.

We plan to investigate a way to learn functions that chronologically order historical photographs in the future.

6. REFERENCES

- [1] “Flickr commons (<http://www.flickr.com/commons>),” .
- [2] Frank Palermo, James Hays, and Alexei A. Efros, “Dating historical color images,” in *ECCV*, 2012.
- [3] Yong Jae Lee, Alexei A. Efros, and Martial Hebert, “Style-aware mid-level representation for discovering visual connections in space and time,” in *ICCV*, December 2013.
- [4] J Severa, *Ordinary Americans & Fashion*, The Kent State University Press, 1997.
- [5] Jayne Shrimpton, *Family Photographs & How to Date Them*, Countryside Books, 2008.
- [6] Neil Storey, *Military Photographs and How to Date Them*, Countryside Books, 2009.
- [7] Jeff Donahue, Yangqing Jia, Oriol Vinyals, Judy Hoffman, Ning Zhang, Eric Tzeng, and Trevor Darrell, “Decaf: A deep convolutional activation feature for generic visual recognition,” *arXiv*, vol. abs/1310.1531, pp. 1–8, 2013.
- [8] Alex Krizhevsky, Ilya Sutskever, and Geoff Hinton, “Imagenet classification with deep convolutional neural networks,” in *NIPS*, 2012.
- [9] Richard W. Haines, *Technicolor Movies: The History of Dye Transfer Printing*, McFarland & company inc. publishers, 1993.
- [10] Grant Schindler and Frank Dellaert, “Probabilistic temporal inference on reconstructed 3d scenes,” in *CVPR*, 2010, pp. 1410–1417.
- [11] Grant Schindler, Frank Dellaert, and Sing Bing Kang, “Inferring temporal order of images from 3d structure,” in *CVPR*, 2007.
- [12] Jiong Xu, Qing Wang, and Jie Yang, “Modeling urban scenes in the spatial-temporal space,” in *ACCV*, 2011.
- [13] Konstantinos Rematas, Basura Fernando, Tatiana Tommasi, and Tinne Tuytelaars, “Does evolution cause a domain shift?,” in *International Workshop on Visual Domain Adaptation and Dataset Bias - ICCV 2013*, 2013.
- [14] A. Ilie and G. Welch, “Ensuring color consistency across multiple cameras,” in *Proceedings of the International Conference on Computer Vision (ICCV)*, October 2005, vol. 2.
- [15] S. J. Kim, H.-T. Lin, Z. Lu, S. Susstrunk, S. Lin, and M. S. Brown, “A new in-camera imaging model for color computer vision and its application,” *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, vol. 34(2), pp. 2289–2302, 2012.
- [16] Y. Xiong, K. Saenko, T. Darrell, and T. Zickler, “From pixels to physics: Probabilistic color de-rendering,” in *Proceedings of the Computer Vision and Pattern Recognition Conference (CVPR)*, Providence, RI, 16/06/2012 2012, IEEE.
- [17] J.J. Koenderink and A.J. Doorn, “Representation of local geometry in the visual system,” *Biological Cybernetics*, vol. 55, no. 6, pp. 367–375, 1987.
- [18] J. van de Weijer, T. Gevers, and A. Gijsenij, “Edge-based color constancy,” *Trans. Img. Proc.*, vol. 16, no. 9, pp. 2207–2214, Sept. 2007.
- [19] A. Gijsenij, T. Gevers, and J. van de Weijer, “Computational color constancy: Survey and experiments,” *Trans. Img. Proc.*, vol. 20, no. 9, pp. 2475–2489, Sept. 2011.
- [20] A. Gijsenij, T. Gevers, and J. van de Weijer, “Generalized gamut mapping using image derivative structures for color constancy,” *International Journal of Computer Vision*, vol. 86, no. 2-3, pp. 127–139, 2010.
- [21] S. Lin, Jinwei Gu, S. Yamazaki, and Heung-Yeung Shum, “Radiometric calibration from a single image,” in *CVPR*, 2004.
- [22] Jianxiong Xiao, J. Hays, K.A. Ehinger, A. Oliva, and A. Torralba, “Sun database: Large-scale scene recognition from abbey to zoo,” in *CVPR*, 2010.
- [23] James Hays and Alexei A Efros, “Im2gps: estimating geographic information from a single image,” in *CVPR*, 2008.